Naval Research Laboratory

Washington, DC 20375-5000

NRL Report 9045

Detecting a Target of Unknown Brightness in Clutter

CHARLES F. OSGOOD AND RICHARD G. PRIEST

Advanced Concepts Branch
Optical Sciences Division

April 27, 1987

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release, distribution unlimited.
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S) NRL Report 9045	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Naval Research Laboratory 6b. OFFICE SYMBOL (If applicable) Code 6520	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Washington, DC 20375-5000	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61153N		RR14-11-4A	DN155-112

11. TITLE (Include Security Classification)

Detecting a Target of Unknown Brightness in Clutter

12 PERSONAL AUTHOR(S)

Osgood, C.F. and Priest, R.G.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Interim	FROM 1/86 TO 9/86	1987 April 27	13

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

Target detection
Algorithms

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

Standard infrared target detection algorithms based on the Bayes decision rule or Neyman-Pearson test are optimal only when testing for targets of known strength (brightness). The discriminant functions used in these tests depend, in functional form, on the assumed brightness of the targets being sought. That is to say, they are not uniform with respect to target brightness. Linear uniform tests such as the matched filter are not near optimal for multidimensional cases. The case of interest here is a multidimensional one—the detection of moving targets in differenced mosaic images. The uniform test that we consider is the generalized maximum likelihood (GML) test. Three implementations are discussed. Results are presented that indicate that the uniform GML test compares favorably with the optimal Bayes decision rule for detection of moving targets in mosaic imagery.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Richard G. Priest	22b. TELEPHONE (Include Area Code) (202) 767-1878 22c. OFFICE SYMBOL Code 6520

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

☆U.S. Government Printing Office: 1985—507-047

CONTENTS

DETECTING A TARGET OF UNKNOWN BRIGHTNESS IN CLUTTER

GENERALIZED MAXIMUM LIKELIHOOD TEST

This effort devises a mathematical technique to permit the detection of a target of unknown brightness that is buried in clutter in a digitized difference frame. The target is idealized to be a point source whose output had been blurred by the atmosphere into a circular Gaussian blur with its center at the position of the target. Therefore, in a difference frame a target should appear as a union of a positive and a negative blur, referred to as a dipole. The dipole is said to be a union of two monopoles. The distance separating the centers of the positive and the negative blurs of the dipole can be bounded above and below by using estimates of maximum and minimum (likely) target speed. It follows that in this model the target does not occlude the clutter. The clutter intensity in each pixel is assumed to be a normally distributed random variable with unknown variance σ that is assumed to be the same in each pixel. Every mean is assumed to be 0. Let target brightness be an unknown parameter denoted by β. Target intensity in each pixel is assumed to be a normally distributed random variable with variance σ where the mean intensity in pixel j is

$$m_j = \frac{\beta}{2\pi\sigma^2} \int\int_{\text{pixel}\, j} \exp\left[\frac{-(u-x_1)^2-(v-x_2)^2}{2\sigma^2}\right] dudv$$

$$- \frac{\beta}{2\pi\sigma^2} \int\int_{\text{pixel}\, j} \exp\left[\frac{-(u-x_3)^2-(v-x_4)^2}{2\sigma^2}\right] dudv\ ,$$

where (x_1, x_2) is the center of the positive blur and (x_3, x_4) is the center of the negative blur. The clutter and target intensities distributions in the different pixels are assumed to be independent random variables.

The assumption is that we have been given F, a digitized difference frame. We need to determine a mathematical test for the presence of such a dipole. The test is to be applied to all 4×4 subframes of F, and it will decide if the target is present. If it decides affirmatively, it will then determine the two positions. The approach uses the generalized maximum likelihood method to obtain a function to be thresholded. Given observations of values of independent random variables each having probability density functions involving an unknown vector of parameters θ, the value of θ that maximizes the joint density function of the random variables (at the values in the reports) converges as the number of reports goes to infinity to the correct value of θ. This is of course, subject to plausible technical conditions. Thus, this maximum value of the joint density function should be a good approximation to the joint density function at the values in the reports. We next deal with maximizations over a subset of the original space of possible observations to test the hypothesis.

Let sets A and B each be complements of the other. In the generalized maximal likelihood test to decide whether to accept the null hypothesis that θ belongs to class A, the maximizations are carried out over the sets A and $A \cup B$, respectively, and one thresholds the ratio of the first to the second. This is an approximation to a normalized likelihood of θ belonging to set A. The value of the threshold is to be between 0 and 1. It is immediate that it suffices to threshold on the ratio of the maximum taken over the class A and the maximum taken over the class B, since if the larger maximum occurs in class B the function to be thresholded is unchanged, while if the larger maximum

Manuscript approved January 7, 1987.

occurs in class A, the new value instead of being 1 (as it would have been) is now greater than 1 and properly exceeds any threshold that is less than or equal to 1.

It suffices to use just one report in the generalized maximum likelihood method if the components are statistically independent and if the dimensionality of the report vector is sufficiently high.[1] Here we have 16 dimensional reports (since each of the 4×4 subframes contains 16 pixels), possibly justifying our use of generalized maximum likelihood. Where $\bar{\theta}$ is the vector $(x_1, x_2, x_3, x_4, \beta)$, let $m(\theta)$ denote the 16 dimensional vector having as its components the m_j. Where $\bar{Y} = (y_i)$ is a vector of observed pixel intensities (for a difference frame), the density function for the jth component is

$$\beta (2\pi)^{\frac{-1}{2}} \sigma^{-1} \exp \left[\frac{-(y_i - m_j(\bar{\theta}))^2}{2\sigma^2} \right] . \tag{2}$$

The two hypotheses are either that a target is present or that a target is not present. If a target is not present, we see that β is 0. In all other cases, a target is present. It follows that in our class B each m_j equals 0, so the density functions are constant there. Thus we must threshold

$$\max_{\theta} \left\{ (2\pi)^{-8} \sigma^{-16} \exp \left[\frac{-\sum_{j=1}^{16} (y_j - m_j(\bar{\theta}))^2 + \sum_{j=1}^{16} (y_j)^2}{2\sigma^2} \right] \right\} . \tag{3}$$

Actually, it sufficies to maximize

$$-\sum_{j=1}^{16} (y_j - m_j(\bar{\theta}))^2 + \sum_{j=1}^{16} (y_j)^2 . \tag{4}$$

Several roads can be followed in attempting to determine the maximum of this expression. The one that seemed to be best for immediately producing numbers is Approach 1 given below. Approach 2 was implemented near the end of this research effort, and only a few numbers were produced. Approach 3 is a more general approach that was not implemented because of lack of time.

APPROACH 1

The idea here is to first maximize Eq. (4) analytically with respect to the brightness β for each set of points p_1 and p_2. Let $\beta\bar{M}$ denote $(m_1(\bar{\theta}), \ldots, m_{16}(\bar{\theta}))$, where $\bar{M}$ is independent of β but is still a function of $p_1 = (x_1, x_2)$ and $p_2 = (x_3, x_4)$. Let $\bar{Y}$ denote $(y_1, \cdots, y_{16})$. Then Eq. (4) equals $-\beta^2 \bar{M} \cdot \bar{M} + 2\beta \bar{Y} \cdot \bar{M}$. Taking the derivative with respect to β, equating it to 0, and solving for β we find that for each choice of p_1 and p_2 the maximum occurs at

$$\beta = \frac{\bar{Y} \cdot \bar{M}}{\bar{M} \cdot \bar{M}} ,$$

and the maximum value is

$$\frac{(\bar{Y} \cdot \bar{M})^2}{\bar{M} \cdot \bar{M}} .$$

[1]Harold Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, N.J., 1974), 13th printing, p. 496.

It therefore suffices to threshold

$$\max_{p_1, p_2} \left\{ \frac{|\bar{Y} \cdot \bar{M}|}{|\bar{M}|} \right\},$$

where $|\bar{M}|$ denotes the norm of $\bar{M}$.

In practice, $|\bar{Y} \cdot \bar{M}| \; |\bar{M}|^{-1}$ is maximized on a set of points chosen so that the maximum value calculated there is very close to the true maximum. Values of p_1 and p_2 where the maximum occurs are then taken to be the target positions. The actual numerical test is carried out as follows: Coordinates are chosen so that the center point of the 4×4 array of pixels is labeled (0, 0) and each pixel has unit length. We assume that $p_1 = (x_1, x_2)$ where $-1/2 \leq x_1 \leq 1/2$ and $-1/2 \leq x_2 \leq 1/2$, while $p_2 = (x_3, x_4)$ where $1/2 \leq |x_3| \leq 3/2$ and $1/2 \leq |x_4| \leq 3/2$. The reason for these constraints is that it is plausible to assume that the subframe chosen has the positive part of the dipole relatively centered and that the target is not so slow as to have the negative part of the dipole in the same pixel as the positive part of the dipole.

Experiments were run to compare the performance of the GML test to that of the Bayes test. In these experiments, the points p_1 and p_2 over which the maximization was carried out were of the form $(j/4, k/4)$, where j and k assume all integral values such that the respective restrictions on the ranges of p_1 and p_2 are satisfied. The experiments were run on an ensemble of Gaussian noise 4×4 frames (background ensemble) and on an ensemble of frames which were the sum of Gaussian noise and an intensity corresponding to a target dipole (target ensemble). The dipole intensities were calculated by using Eq. (1) with the points p_1 and p_2 chosen randomly from their respective ranges. A false alarm is counted when the test identifies a member of the background ensemble as a target; a missed detection is counted when the test identifies a member of the target ensemble as noise. The important results of these experiments are shown in Figs. 1, 2, and 3. In these figures, the threshold value does not appear explicitly. Rather, the percentage of false alarms that results when the threshold is set to produce a specified number of missed detection is presented. Figure 1 presents the results of the GML test described above. The signal-to-noise ratio of the dipole targets to the Gaussian noise is indicated in the figure legend. Figure 2 presents a comparison of the GML test to the Bayes test that is optimal for the strengths of the targets actually used in the experiment (S/N = 3.92). The GML test performs only slightly less well than the optimal test. Figure 3 compares the performance of the GML test to that of a Bayes test that is not optimal for the target strengths of the target ensemble used, but rather is optimal for much stronger ($10\times$) targets. In this case, the GML test is superior to the Bayes test for low false alarm rates. This result illustrates the value of a test that is invariant to target strength.

APPROACH 2

In our second approach, we searched independently for the positive and the negative monopoles of the dipole. Adopting this search procedure meant deciding that the two optical blurs do not cancel each other out to any appreciable extent. This seems a reasonable assumption when the parameter σ in the optical blur is small compared to the expected distance between the positive and negative monopoles of the dipole. The test has the hypothesis that two monopoles exist and that they are located in the respective regions indicated in Approach 1. The alternative hypothesis is that either 0 or 1 monopoles of either sign exist. (More than two monopoles are regarded as too unlikely an event to be considered. Detecting two monopoles that have the same sign is also a possibility; it is assumed that this case also is sufficiently unlikely that it can be ignored.) Detecting both a positive and a negative monopole but with the regions reversed is allowed. Such occurrences are likely even if all possible 4×4 subframes of a frame are tested, but they should cause little harm since examination would reveal which monopole was which. We assume that the three possible cases, 0 monopoles, 1 monopole, or 2 monopoles have equal prior probabilities of occurrence. We use generalized maximal likelihood in formulating our test and assumed that the brightnesses of the two monopoles are independent. This is a somewhat unrealistic assumption, although it does somewhat compensate for the effect

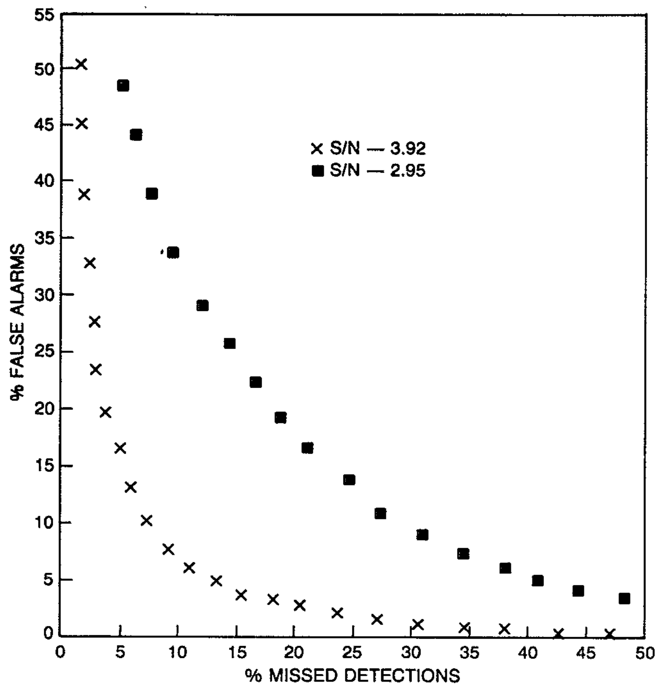

Figure 1 — GML test

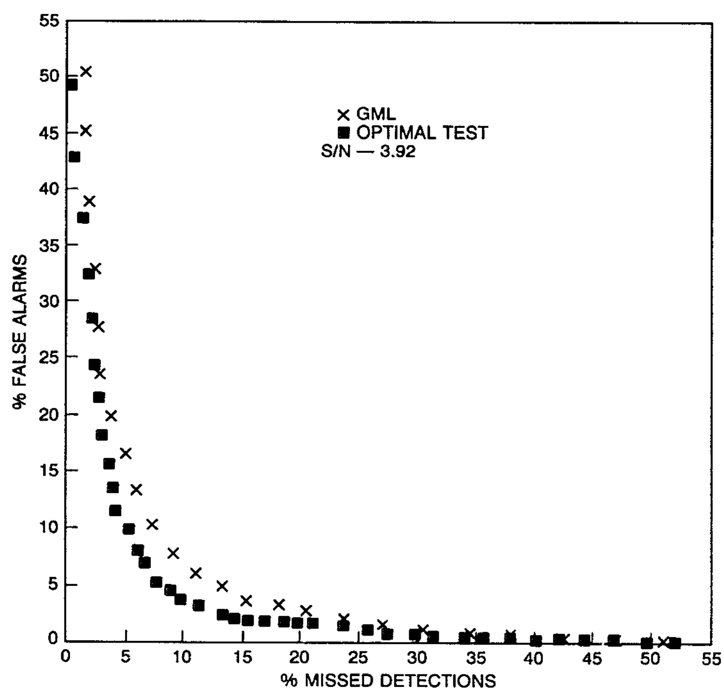

Figure 2 — GML test vs optimal test

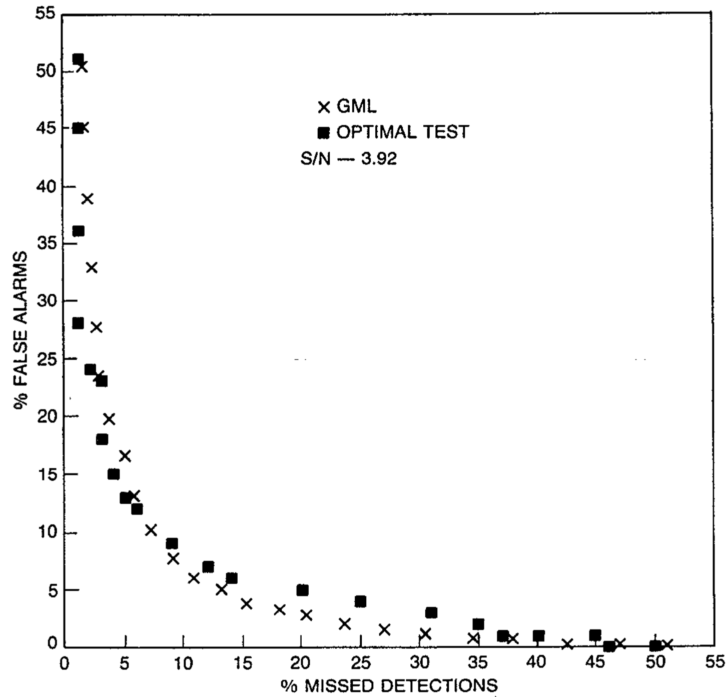

Figure 3 — GML test vs test optimal for strong targets

of any cancellation between the positive and negative monopoles, as well as any cancellation due to clutter. Below, let $m_j(p)$ for $j = 1, 2, \ldots, 16$, denote the mean brightness of pixel j under the assumptions of a positive monopole of unit brightness at p. Let β_1 and β_2 denote the respective brightnesses of the two monopoles, and let $p_1 = (x_1, x_2)$ and $p_2 = (x_3, x_4)$ denote the two monopole positions.

We must evaluate

$$\Gamma = \max_{\beta_1, \beta_2, p_1, p_2} \left\{ (2\pi)^{-8} \sigma^{-16} \exp \left(\frac{-\sum_{j=1}^{16} (y_j - \beta_1 m_j(p_1))^2 - \sum_{j=1}^{16} (y_j - \beta_2 m_j(p_2))^2}{2\sigma^2} \right) \right\}$$

divided by

$$\frac{1}{3} \max_{\beta_1, \beta_2, p_1, p_2} (2\pi)^{-8} \sigma^{-16} \left\{ \exp \left(\frac{-\sum_{j=1}^{16} (y_j - \beta_1 m_j(p_1))^2 - \sum_{j=1}^{16} (y_j)^2}{2\sigma^2} \right) \right.$$

$$+ \exp \left(\frac{-\sum_{j=1}^{16} (y_j - \beta_2 m_j(p_2))^2 - \sum_{j=1}^{16} (y_j)^2}{2\sigma^2} \right)$$

$$+ \exp \left[\frac{-2 \sum_{j=1}^{16} (y_j)^2}{2\sigma^2} \right] \Bigg\} .$$

Our strategy is to maximize analytically with respect to β_1 and β_2, in both the numerator and the denominator, and carry out the remainder of each maximization by exhaustive search. The two βs always occur in distinct quadratics, i.e., in the

$$-(2\sigma)^{-1} \sum_{j=1}^{16} (y_j - m_j\ (\beta_i, p_i))^2 ,$$

so it suffices to maximize each independently with respect to β_i. Such a calculation was carried out above in complete detail when we were testing simultaneously for both monopoles of a dipole.

Let $\bar{Y}$ denote (y_j). Where $\bar{M}(p_i)$ is the vector $(m_j(p))$, each of the above-mentioned quadratics has a maximal value

$$(2\sigma^2)^{-1} \max_{p_i \epsilon R_i} \left\{ \frac{(\bar{Y} \cdot \bar{M}(p_i))^2}{|\bar{M}|^2} \right\} - (2\sigma^2)^{-1} |\bar{Y}|^2 ,$$

where R_i is the region containing p_i.

Dividing the numerator and the denominator of the expression defining T by

$$\exp\ (-(\sigma^2)^{-1} |\bar{Y}|^2) ,$$

it follows that for $i = 1, 2$, where

$$T_i = (2\sigma^2)^{-1} \max_{p_i \epsilon R_i} \left\{ \frac{(\bar{Y} \cdot \bar{M}(p_i))^2}{|\bar{M}|^2} \right\} ,$$

$$3T = (\exp\ (T_1 + T_2))\ (1 + \exp(T_1) + \exp\ (T_2))^{-1} .$$

For large T_1 and T_2, one may threshold by using the function $\min\{T_1, T_2\}$. It can be seen that this is correct by dividing both numerator and denominator by the larger of $\exp\ (T_1)$ and $\exp\ (T_2)$ and taking logarithms.

APPROACH 3

Here we adhere to the convention that an overbar denotes a vector, and a lower-case letter with a subscript denotes a vector component.

This approach is a different version of Approach 1. Suppose that we do not want to use an exhaustive search to maximize Eq. (4). This could become intractable, yet an analytic approach seems to be too complicated to carry out. A promising simplification is to replace Eq. (1) in each pixel by a linear function of the two target positions. To be definite, let our linear function be the linear polynomial in two variables that, together with its first partial derivatives, agrees with the function in Eq. (1) at the center of the pixel. Therefore, the $m_j(\theta)$ are linear polynomials in four variables. Set

$$m_j(\overline{\theta}) = \sum_{k=1}^{4} a_{jk} z_k + a_j , \tag{5}$$

or equivalently,

$$\overline{M} = \sum_{k=1}^{4} z_k \overline{a}_k + \overline{a} , \tag{6}$$

$$\overline{a}_k = (a_{jk}) \text{ for } k = 1,2,3,4, \text{ and } \overline{a} = (a_j).$$

By the Gram-Schmidt orthogonalization procedure, we can produce from the sequence of vectors $\overline{a}_1$, $\overline{a}_2$, $\overline{a}_3$, $\overline{a}_4$, and $\overline{a}$ a sequence of orthogonal vectors $\overline{\alpha}_1$, $\overline{\alpha}_2$, $\overline{\alpha}_3$, $\overline{\alpha}_4$, and $\overline{\alpha}$ such that real numbers c_{kj} and c_j exist that satisfy

$$\overline{a}_k = \sum_{j \le k} c_{kj} \overline{\alpha}_j ,$$

$$\overline{a} = \sum_{j=1}^{4} c_j \overline{\alpha}_j + \overline{\alpha},$$

$$|\overline{\alpha}_1| = \cdots = |\overline{\alpha}_4| = |\overline{\alpha}| ,$$

$$\begin{aligned} \overline{M} &= \sum_{k=1}^{4} z_k \left[\sum_{j \le k} c_{kj} \overline{\alpha}_j \right] + \sum_{j=1}^{4} c_j \overline{\alpha}_j + \overline{\alpha} \\ &= \left[\sum_{j=1}^{4} \left[\sum_{k=1}^{4} c_{kj} z_k + c_j \right] \overline{\alpha}_j \right] + \overline{\alpha} \\ &= \sum_{j=1}^{4} w_j \overline{\alpha}_j + \overline{\alpha}, \end{aligned}$$

$$w_j = \left[\sum_{k=1}^{4} c_{kj} z_k + c_j \right] ,$$

Set $\bar{Y} = (y_j)$, and recall that

$$\bar{M} = (m_j) = \sum_{j=1}^{4} w_j \bar{\alpha}_j + \bar{\alpha}.$$

Then the function to be minimized is

$$-2\beta \left(\sum_{j=1}^{4} w_j (\bar{\alpha}_j \cdot \bar{Y}) + (\bar{\alpha} \cdot \bar{Y}) \right) + \beta^2 \left(\sum_{j=1}^{4} |\bar{\alpha}|^2 (w_j)^2 + |\bar{\alpha}|^2 \right)$$

$$= |\bar{\alpha}|^2 \left[\sum_{j=1}^{4} \left(\beta w_j - \frac{\bar{\alpha}_j \cdot \bar{Y}}{|\bar{\alpha}|^2} \right)^2 + \left(\beta - \frac{\bar{\alpha} \cdot \bar{Y}}{|\bar{\alpha}|^2} \right)^2 \right.$$

$$\left. - \left(\sum_{j=1}^{4} \frac{(\bar{\alpha}_j \cdot \bar{Y})^2}{|\bar{\alpha}|^2} + \left(\frac{(\bar{\alpha} \cdot \bar{Y})^2}{|\bar{\alpha}|^2} \right) \right) \right].$$

The first term inside the square brackets is $|\bar{\alpha}|^2$ times the square of the distance from the point

$$\bar{p} = \left(\frac{\bar{\alpha}_1 \cdot \bar{Y}}{|\bar{\alpha}|^2}, \frac{\bar{\alpha}_2 \cdot \bar{Y}}{|\bar{\alpha}|^2}, \frac{\bar{\alpha}_3 \cdot \bar{Y}}{|\bar{\alpha}|^2}, \frac{\bar{\alpha}_4 \cdot \bar{Y}}{|\bar{\alpha}|^2}, \frac{\bar{\alpha} \cdot \bar{Y}}{|\bar{\alpha}|^2} \right)$$

to the nearest point of the form

$$(\beta w_1, \beta w_2, \beta w_3, \beta w_4, \beta)$$

where $-\infty < \beta < +\infty$, each w_j is defined as above, and every z_j lies between $-1/2$ and $1/2$. Call this region R.

If $\bar{p}$ belongs to R, the distance is 0, and we have our minimum. The region R is the projection through the origin in 5 space of a 4-space figure H that is the image of the unit hypercube $-1/2 \le z_j \le 1/2$ under the transformation defining the w_js. Call this transformation T. Since one could write the $\bar{\alpha}_j$ and $\bar{\alpha}$ in terms of the $\bar{a}_k$ and $\bar{a}$ and go back through the argument above, T is invertible. Equations of the form $\bar{u} \cdot (z_1, \cdots, z_4) = C_i$, where $\bar{u}$ is a vector of constants and the C_i are a set of constants, are transformed by T into equations of the form $\bar{v} \cdot (z_1, \cdots, z_4) = C_i$ where $\bar{u}$ and T determine $\bar{v}$. Thus parallel planes are taken into parallel planes. Furthermore, our mapping preserves functional values: the new linear function has the same value at the image under T of a point $\bar{q}$ as the original linear function had at the point $\bar{q}$. Points $\bar{q}$ of the unit hypercube are characterized as being simultaneously between four pairs of hyperplanes. (We call the most general solid defined in this manner a hyperparallelipiped.) Four pairs of linear functions exist that vanish on these planes such that for each pair of functions: every $\bar{q}$ in the unit hypercube either takes on nonzero values of opposite sign or exactly one function equals zero at $\bar{q}$. The only points $\bar{q}$ satisfying these conditions for all four pairs of hyperplanes are the points of the unit hypercube. Therefore H, the image of the unit hypercube, is a hyperparallelipiped.

All of four-dimensional space lies in the projections from $T(\bar{0})$ of the faces of H. Let $\bar{p}_1$ be a point of four-dimensional space lying outside of H. It is geometrically plausible that a face of H is a closest face to $\bar{p}_1$ (meaning that the minimum of the distance from $\bar{p}_1$ to a point of H is attained at a point in this face) if and only if $\bar{p}_1$ is in the projection of this face from $T(\bar{0})$. Each of these eight (closed) regions is bounded by planes Π_l through $T(\bar{0})$ and through the hyperedges of the hyperface.

As we have seen, the property of being enclosed by hyperplanes is preserved under the mapping T, and, using the same argument, by the mapping T^{-1}. Apply T^{-1} to $\bar{p}_1$ and test which faces of the unit hypercube are closest to $T^{-1}(\bar{p}_1)$. The faces of the unit hypercube satisfy the equations

$$y_i = \frac{1}{2}$$

or

$$y_i = \frac{-1}{2}.$$

Since the sides of the unit hypercube are orthogonal, the closest point of the unit hypercube to $T^{-1}(\bar{p}_1)$ has each coordinate determined by the requirement that it be the closest point in the interval $[-1/2, 1/2]$ to the corresponding coordinate of $T^{-1}(\bar{p}_1)$. Thus the closest faces can be read off, and the corresponding faces are the closest faces of H to $\bar{p}_1$. However, we are after a closest point in five-dimensional space so we are not yet finished.

In five-dimensional space, R is the projection of H through the origin, and each of the eight faces of H are projected into five-dimensional hyperplanes through the origin. These latter hyperplanes form the boundary of R. Clearly if β is positive (negative), the nearest point of R to $\bar{p}$ must be in the half plane where β is positive (negative). Without loss of generality, assume that β is positive. The projections of the Π_l are five-dimensional hyperplanes through the point $(T(0), 1)$, using obvious notation, and the hyperedges of R. It is geometrically plausible that the closest faces of the cone to a point $\bar{p} = (\beta w_1, \cdots, \beta w_4, \beta)$ are the faces where the point $\bar{p}$ is in the closed region bounded by the hyperplanes through the boundaries of each face. Then it follows that the closest faces of R to $\bar{p}$ are the projections of the closest faces of H to $(w_1, \cdots, w_4)$.

One can determine the normals to each of the eight faces of the five-dimensional cone. For this set of closest faces of the cone, we may use the Gram-Schmidt normalization procedure to produce an orthonormal basis $\bar{e}_1, \cdots, \bar{e}_l$. Then,

$$\bar{p} - \sum_j (\bar{p} \cdot \bar{e}_j)\, \bar{e}_j$$

is the projection on these faces simultaneously so

$$\sum_{j=1}^{l} (\bar{p} \cdot \bar{e}_j)^2$$

is the required squared distance.